AMCIS 2020 Proceedings

AI and Semantic Technologies for Intelligent Information Systems (SIGODIS)

Aug 10th, 12:00 AM

Investigating Data Visualization Decision Errors: Do Tools Enable Users to Make Bad Decisions and more confidently?

Jim Jenkins
University of North Carolina at Greensboro, jhjenkin@uncg.edu

Rahul Singh
UNC Greensboro, rahul@uncg.edu

Investigating Data Visualization Decision Errors: Do Tools Enable Users to Make Bad Decisions and more confidently?

Emergent Research Forum (ERF)

Jim Jenkins
University of North Carolina at Greensboro
jhjenkin@uncg.edu

Rahul Singh
University of North Carolina at Greensboro
rahul@uncg.edu

Abstract

The current generation of Analytics and Business Intelligence (ABI) tools are easy to use and present relevant information in an elegant form. Their promise to "bring analytics to everyone in the company" would mean that the users creating visualizations and making decisions based on these visualizations may not have the training in data science to understand if they are drawing conclusions that are supported by the data. This study proposes to investigate the impact of cognitive biases and the Kruger-Dunning effect within the use of ABI tools on the resultant decision quality and the confidence the decision maker places in that decision.

Keywords

Analytics, business intelligence, cognitive bias, data visualization, Kruger-Dunning effect.

Introduction

Contemporary Analytics and Business Intelligence (ABI) tools are easy to use and present relevant information in an elegant form. ABI tools present complex problems in a simplified manner, allowing users to extract organizational data and curate them for visual exploration. They bring context, relevant information, and models based on complex descriptive and predictive analytics-based algorithms to the end user and invite users to explore data visually, through flexible and model-driven analysis and visualizations. Visual data exploration has particularly high value for users engaged in solving novel and unstructured decision problem where exploring the context is especially useful (Keim, 2002). We refer to this process as Visual data exploration using ABI tools. Like strategic decision making, Visual data exploration processes are ambiguous and have no pre-defined success criteria (Levine, 1971). Research around the simplification of strategic decisions shows that the process of simplifying the problem within the decision process may lead to the introduction of bias (Gigerenzer and Gaissmaier, 2011; Schwenk, 1984). Yet, ABI tools are increasingly adopted by users to make business decisions that have significant impact on organizations and individuals.

The ABI market is vast, rapidly growing and includes a wide range of tools. Forbes recently reported that the ABI market in 2016 was worth $130 billion and growing at a rapid 11.7% CAGR, with projections reaching $203 billion of worldwide spend by the end of that year. The 2018 Gartner Magic Quadrant Analysis for Analytics and Business Intelligence (ABI) tools includes 20 commercial product offerings (Howson, et al., 2018). With the increasing rate at which data is being accumulated, the shortage of trained and experienced data scientists, the ease at which data visualizations can be created using the latest generation of data visualization tools and the promise of these data visualization tools to "bring analytics to everyone in the company" (Qlick website). In the US, demand for data scientists far exceeds the number of trained data scientists available, a shortage of over 150,000 data scientists was reported in August, 2018 (LinkedIn, 2018). This has fueled the growth of citizen data scientists whose primary job function and training is outside the fields of statistics or analytics. These citizen data scientists build and use data visualizations, predictive and descriptive models of data and influence or make decisions that have significant impact on organizations. Gartner projects that the number citizen data scientists will increase at five times the rate of expert data scientists (Howson et al. 2018).

This paper investigates the research question: *Does the wide-spread and phenomenal growth in use of ABI tools for analysis and decision-making lead to rapid decision making, with high user-confidence, while compromising decision quality through the systematic introduction of cognitive biases?*

As Big Data and ABI tools with data visualization become more widely used (Henke et al, 2016), the potential impact of cognitive bias in decision making will continue to grow. It is important to consider the impact of long-standing research on the manifestation of cognitive biases in ABI tools and data visualization. Without this knowledge, the impact of biases will remain hidden and decision quality will suffer. This paper investigates the opportunities for bias to impact the decision processes enhanced by ABI tools. Understanding the sources and potential impacts of cognitive biases can inform practitioners and data scientists as they create their models and decision aids to reduce the effects of bias to increase the quality of the decisions made using these tools.

Theoretical Foundations

Studying statistical intuitions of experts, Tversky and Kahneman (1971) found excessive confidence and the persistence of systematic errors in the intuitions of experts. Intuitive judgments are often incorrect and ignore the more deliberate computations experts are trained to execute (Kahneman and Fredrick, 2002). Moreover, greater expertise leads to stronger perceptions of efficacy. Tversky and Kahneman classified ways that decision makers systematically diverge from rational actions as forms of Cognitive Bias (Kahneman & Tversky, 1972; Tversky & Kahneman, 1974). When the lines between intuitive judgements, perceptions and deeper systematic thinking that leads to accurate decisions are blurred, there is greater confidence in the decision. When users are under-trained in the tools or business context, they are less likely to be aware of their biases, resulting in over-confidence in decision quality (Kruger and Dunning, 1999). Although they have been long considered in unaided decision-making processes (Campitelli & Gobet, 2010; Tversky & Kahneman, 1974, 1981, 1989), the role and potential impact of biases in decision making with ABI tools and visual exploration have received little attention in the IS literature .

Cognitive Bias and Decisions

Considering only the human component of a decision, Evans (1989) proposed two distinct stages of a decision process, which he called a "heuristic stage" and an "analytic stage". Evans used "heuristic" to refer to the pre-attentive process wherein aspects of the problem are curated for their relevance. This process determines what the decision maker will attend to and think about as they consider the decision (Evans, 1989). We will adapt and extend Evans' model, presented in Figure 1, to guide this research.

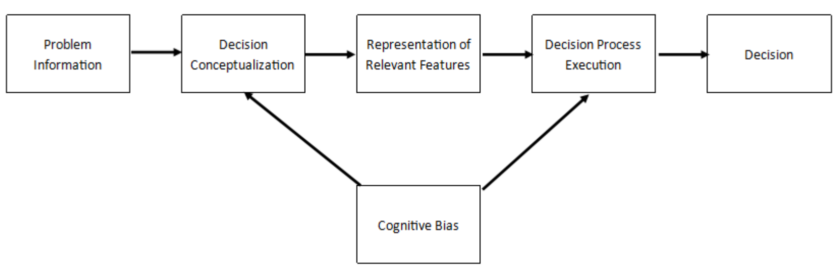

Figure 1. Conceptual Model

In conceptualizing decisions, the decision-maker determines which aspects of the information are important, and what process they will use to make the ultimate decision. Evans (1989) discussed factors such as attentional salience of presentations, linguistic factors and the effects of prior knowledge on this process. These influence the representation of the relevant features of the problem. The representation has significant influence on the decision and is subject to cognitive bias, often without the decision maker being aware of it. Arnott and Gao (2019) note that as this research stream is "a large collection of interrelated theories where each theory describes a particular aspect of human decision-making, it is difficult to have a coherent overall view of the research". We summarize the research on biases relevant for this study below.

Decision makers using ABI tools expose themselves and the decisions that they make based on these visualizations to several different forms of cognitive bias. In their review of the use of cognitive biases in general IS research, Fleischmann et al. (2014) showed that framing bias and anchoring bias were the most popular biases considered in IS research. The data visualization tool potentially empowers decision makers to make judgements that they may not be equipped to make. Moreover, the decision makers have lower awareness and greater confidence in their decisions, irrespective of the decision quality, indicating that decision makers are biased partly due to their lack of awareness of, or indifference towards, their biases (Pennycook at al 2017).

Research Model and Hypotheses.

We study the impact of systematic cognitive biases on decision quality, in decision making processes facilitated through ABI and visual data exploration tools. Following is our research model and hypothesis.

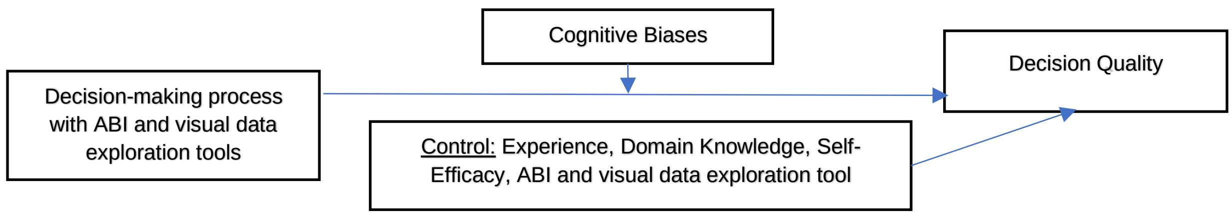

Figure 2. Research Model

Attenuation bias refers to over-simplifying a decision situation by ignoring or significantly discounting the level of uncertainty (Arnott, 1998). When exhibiting attenuation bias, a decision maker may reduce the complexity of decisions in uncertain environments by simply declaring that it is certain (Remus and Kottemann, 1986), often by discarding or discounting information or decision conceptualizations that contribute to the uncertainty (Hogarth, 1987). In visual data exploration using ABI tools, attenuation bias affects decision outcomes when the user curates and organizes the data based on arbitrary and inconsistent criteria. These criteria are selected for congruence to a narrative or ease of analysis and explanation, rather than for their ability to completely and accurately describe a phenomenon. We hypothesize that:

H1: Variability will be reduced or ignored to simplify the decision and differences between two series will be seen more often than they exist.

Automation bias or automation-induced complacency can manifest as either *errors of omission* (the user misses events because the system did not require action) or *errors of commission* (the user follows the recommendation of the automation, even when this recommendation contradicts their training and other available data)(Parasuraman & Manzey, 2010).

H2: Users follow the recommendation of the system, even when they come to different conclusions upon reflection.

Clustering Illusion or Chance bias is the tendency to under-estimate the consequence of variability and infer implications of patterns from small sets of random data (Tversky & Kahneman, 1971). This bias can be exemplified by asking a person: *what are the odds of a coin flip resulting in a heads outcome after ten consecutive tails have been flipped?* Humans as a group are particularly poor at perceiving randomness (Tversky & Kahneman, 1974; Hogarth, 1987). In visual data exploration using ABI tools, this leads to users being more certain of causal inferences in data and their general implications, even when the data is a small fraction of the population.

H3: Users infer more conclusions than are supportable by the data and have higher degree of confidence in the conclusions that they draw from visualizations

Confirmation Bias manifests as a tendency to search for or interpret evidence that confirms *a priori* conclusions, while not searching for disconfirming information. Confirmation bias is a well-known and "widely accepted notion of inferential error to come out of the literature on human reasoning" (Evans, 1989). Confirmation bias leads to undue bolstering of beliefs whose truth is in question (Nickerson, 1998).

H4: Users express higher confidence for conclusions that agree with a priori expectations, with little supporting evidence.

Framing Bias is the tendency to evaluate events differently depending on how they are presented (Tversky & Kahneman, 1981). For example, when presented with treatment descriptions described in positive, negative, or neutral terms, adults are significantly more likely to agree to a treatment when it is positively described than they are to agree to the same treatment when it is described neutrally or negatively (Peters et al, 2000). In visual data exploration using ABI tools, the presentation of a problem domain and data selection for exploration creates multiple opportunities for framing bias.

H5: Interpretations of visualizations are influenced by framing of requested interpretation.

Mode Bias occurs because the manner in which data is presented can affect how it is interpreted (Hogarth, 1987). A series of studies in the 1970s investigating the impact of data presentation to managers showed that when the information was presented in a manner that was more engaging to managers, the more they valued data (Bhappu, Griffith, & Northcraft, 1997; McKenney & Keen, 1974). In visual data exploration using ABI tools, more engaging and "polished" data presentations are likely to be assigned greater value.

H6: Users place higher confidence in conclusions drawn from data presented in more visually appealing graphics types.

Overconfidence Bias is the tendency of the decision maker to overestimate their ability to solve difficult or novel problems (Brenner et al, 1996). The amount of domain knowledge a person has greatly affects the magnitude and the direction of the effect of this bias. Users with lower domain experience are more likely to be over-confident of their conclusions while users with more domain experience are more likely to report lower confidence in their conclusions (Brenner et al, 1996). In visual data exploration using ABI tools, we expect *citizen data scientists* to report higher confidence in the conclusions than domain experts, irrespective of decision quality.

H7: Users with lower domain experience express higher confidence in conclusions drawn from visualizations than what the data support.

Sample Bias is the tendency to ignore sample size when determining predictive power (Tversky & Kahneman, 1971). The "Law of Large Numbers" holds that relatively large samples will be highly representative of the population they are drawn from (Newey & McFadden, 1994). The Sample bias arises when this law is thought to hold for small samples. The Sample bias has been found to hold even for academic psychologists who specialize in statistics when they design experiments or consider changes to non-significant experiments. (Tversky & Kahneman, 1974; Sedlmeier & Gigerenzer, 1997).

H8: Confidence in conclusions drawn from visualizations exceeds what the data support.

Method

An experimental design will be used to test our hypothesis and draw conclusions from the research model. We will test individuals across different levels of subject matter expertise and exposure to statistical methods. These two dimensions will be self-reported by the subjects and will be used to group the subjects into four categories. These categories will represent hi and low values across each dimension. Each subject will be given a series of AVI representations and asked to select from a series of conclusions that they might draw from these representations of the data and to indicate their confidence in their selection. The conclusions that each subject chooses will be compared to a known set of best choices. The results for each of the four groups will be aggregated and tested using an ANOVA with a Tukey-Kramer correction for pairwise comparisons.

Conclusion and Future Research

This research contributes to the information systems literature concerning analytics and decision making by proposing validated theoretical foundations to understand the impact of nascent cognitive biases introduced and manifest in the process of decision making with ABI and visual data exploration platforms. This research also informs practitioners so they can be more cognizant of how their cognitive biases could influence the decisions that they are augmenting with machine learning and data visualizations.

REFERENCES

Arnott, David. 1998. A Taxonomy of Decision Biases. In *A Taxonomy of Decision Biases*.

Arnott, D., & Gao, S. 2019. Behavioral economics for decision support systems researchers. *Decision Support Systems*, (*122*), 113063. (https://doi.org/10.1016/j.dss.2019.05.003)

Bhappu, A.D., Griffth, T.L., & Northcraft, G.B. 1997. Media effects and communication bias in diverse groups. Organisational Behaviour and Human Decision Processes, (70:3), 199-205.

Brenner, L.A., Koehler, D.J., Liberman, V., & Tversky, A. 1996. Overconfidence in probability and frequency judgements: A critical examination. *Organisational Behaviour and Human Decision Processes*, (65), pp. 212-219.

Campitelli, G., & Gobet, F. 2010. Herbert Simon's Decision-Making Approach: Investigation of Cognitive Processes in Experts. *Review of General Psychology*, (*14*:4), pp. 354–364.

Evans J.St.B.T. (1989). Bias in human reasoning: causes and consequences. Lawrence Erlbaum.

Fleischmann, M., Amirpur, M., Benlian, A., & Hess, T. 2014. COGNITIVE BIASES IN INFORMATION SYSTEMS RESEARCH: A SCIENTOMETRIC ANALYSIS. *European Conference on Information Systems (ECIS)*, pp. 0–21.

Gigerenzer, G., & Gaissmaier, W. 2011. Heuristic decision making. *Annl review of psych.*, *62*, 451-482.

Henke, N., Bughin, J., Chui, M., Manyika, J., Saleh, T., Wiseman, B., & Sethupathy, G. 2016. *The Age Of Analytics: Competing In A Data-Driven World*. McKinsey Global Institute.

Hogarth, R. (1987). Judgement and choice: The psychology of decision (2nd ed.). Chichester, UK: Wiley

Howson, C., Sallam, R., Richardson, J., Tapadinhas, J., Idoine, C., & Woodward, A. (2018, February 26). Magic Quadrant for Analytics and Business Intelligence Platforms (Tech. No. G00326555).

Kahneman, D., & Frederick, S. 2002. Representativeness revisited: Attribute substitution in intuitive judgment. *Heuristics and biases: The psychology of intuitive judgment*, (*49)*, pg. 81.

Kahneman, D., & Tversky, A. 1972. Subjective probability: A judgment of representativeness. *Cognitive Psychology*, (*3*:3), pp. 430–454.

LinkedIn Economic Graph Team. 2018. LinkedIn Workforce Report - United States. August 2018.

Keim, D. A. 2002. Information visualization and visual data mining. *IEEE transactions on Visualization and Computer Graphics*, (*8*:1), pp. 1-8.

Kruger, J., & Dunning, D. 1999. Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *J. of Personality and Social Psychology, (77),* pp. 1121–1134.

Levine, M. 1971. Hypothesis theory and nonlearning despite ideal SR-reinforcement contingencies. *Psychological review*, (*78*:2), pg. 130.

McKenney, J., & Keen, P. 1974. How managers' minds work. Harvard Business Review, (72:3), pp. 79-90.

Newey, K. W., & McFadden, D. 1994. Large sample estimation and hypothesis. *Handbook of Econometrics, IV,* Edited by RF Engle and DL McFadden, pp. 2112-2245.

Nickerson, R. S. 1998. Confirmation bias: A ubiquitous phenomenon in many guises. Review of general psychology, (2:2), pp. 175-220.

Parasuraman, R., & Manzey, D. 2010. Complacency and Bias in Human Use of Automation: An Attentional Integration. *Human Factors*, (*52*:3), pp. 381–410.

Peters, E., Finucane, M., MacGregor, D., & Slovic, P. 2000. The bearable lightness of aging: Judgment and decision processes in older adults. *The aging mind: Opportunities in cognitive research*, pp. 144-165.

Remus, W.E., & Kottemann, J.E. (1986). Toward intelligent decision support systems: An artificially intelligent statistician. MIS Quarterly, (10:4), pp. 403-418.

Schwenk, C. R. (1984). Cognitive simplification processes in strategic decision-making. *Strategic management journal*, (*5*:2), pp. 111-128.

Tversky, A., & Kahneman, D. 1971. Belief in the law of small numbers. *Psychological bulletin*, (*76*:2), pg. 105.

Tversky, A., & Kahneman, D. 1974. Judgment under Uncertainty : Heuristics and Biases. *Science*, (*185*:4157), pp. 1124–1131. (https://doi.org/10.1126/science.185.4157.1124)

Tversky, A., & Kahneman, D. 1981. The Framing of Decisions and the Psychology of Choice. *Science*, (*211*:4481), pp. 453–458.

Tversky, A., & Kahneman, D. 1989. Rational Choice and the Framing of Decisions. In *Multiple criteria decision making and risk analysis using microcomputers*. Berlin: Springer. .pp. 81–126.